# Kapitel VII.

## Differentialrechnung.

Von *Paul Epstein* in Straßburg i. E.

---

### Einleitung.

Seit den ersten Jahrzehnten des 17. Jahrhunderts vollzieht sich die Erweiterung des mathematischen Gedankenkreises (*infinitesimale Betrachtungen, veränderliche Größen*), die schließlich in der Schöpfung der Differential- und Integralrechnung gipfeln sollte. Es waren in erster Linie *geometrische* Probleme, an denen sich die Fruchtbarkeit der infinitesimalen Betrachtungen bewährte: zunächst — bewußt anknüpfend an *Archimedes* — *Flächen-* und *Körperberechnungen*, dann das *direkte* und *inverse Tangentenproblem, Maxima* und *Minima, Rektifikationen.* Daneben entwickelt sich, vielfach durch *mechanische* Vorstellungen beeinflußt, der Begriff (oder besser das Gefühl) der *stetigen Veränderung.* Die folgende Übersicht mag die angedeutete Entwicklung in ihren Hauptzügen kennzeichnen. Wegen aller Einzelheiten, insbesondere wegen des Zusammenhangs mit der Philosophie des Mittelalters (Bradwardinus, Oresme, Nicolaus von Cues) muß auf Cantor, *Gesch. d. Math.* **2** und **3** verwiesen werden. Vgl. ferner Zeuthen, *Gesch. d. Math. im 16. u. 17. Jahrh.*

Kepler (1571—1630), *Stereometria doliorum* 1615, Flächen- und Körperberechnungen, Maxima und Minima. Cavalieri, *Geometria indivisibilibus continuorum . . . promota* 1635, *Exercitationes geometricae.* Flächen- und Körperberechnungen, Schwerpunktsbestimmungen. Descartes (1596—1650), *Geometrie* 1637. *Briefwechsel*, Tangentenproblem (vgl. Tannery, 3. intern. Math.-Kongr. 502 (1904)), Rektifikationen, Kurvenuntersuchungen. Fermat (1601—1665, der bedeutendste Vorgänger von Newton und Leibniz), Maxima und Minima (zwischen 1629 und 1638, veröffentl. 1679 in *Varia Opera*, enthält den Begriff des Differentialquotienten), Tangentenproblem (veröffentl. 1642), Flächen-

berechnungen (*Varia Opera* 44), Rektifikationen (*de linearum curvarum cum lineis rectis comparatione* 1660). Toricelli, *Opera Geometrica* 1644. Roberval, *Observations sur la composition des mouvements* (1668, aber schon 1644 erwähnt), Anwendung des Kräfteparallelogramms auf das Tangentenproblem. Wallis, *Arithmetica infinitorum* 1655, Potenzsummen, Grenzübergang. Pascal (1623—1662), *Œuvres* **3** (1872) Potenzsummen, Partielle Integration, Charakteristisches Dreieck. Huygens (1629 —1695), *Horologium oscillatorium* 1673 (vollendet 1665). Evoluten und Evolventen. Krümmungsmittelpunkt. Maxima und Minima, Tangentenproblem (zwischen 1652 und 1667, sehr nahe übereinstimmend mit *de Sluse* 1652). Barrow (vgl. Zeuthen, 1. intern. Math. Kongr. 274 (1897)), *Lectiones opticae et geometricae* 1669. Tangentenproblem auf Grund der Bewegungslehre. Newton (1643—1727), *Analysis per aequationes* 1666. Flächen- und Körperberechnungen und Rektifikationen als gleichartige Aufgaben erkannt. *Methodus fluxionum* (1671, veröffentl. 1736). Die Bezeichnungen *Fluxion* (Differentialquotient) und *Fluente* zum ersten Mal in einem *Brief an Leibniz* vom 24. 10. 1676. Leibniz (1646—1716) (Leibniz' math. Schriften, herausg. von Gerhardt 1849—1863, Gerhardt, Die Entdeckung d. höh. Analysis 1855). Briefwechsel und Manuskripte von 1670 —1675. Ausbildung der *Bezeichnung*. *Endgültige Lösung des Tangentenproblems* in der Antwort auf Newtons Brief.

Der große Fortschritt von Newton und Leibniz besteht in der Schöpfung eines *Algorithmus*, der die große Menge der vorausgehenden Einzeluntersuchungen unter einheitlichen Gesichtspunkten zusammenfaßt. Wem von beiden das Hauptverdienst zukommt, kann dahingestellt bleiben; sicher war Newton vor Leibniz im Besitz wichtiger Erkenntnisse, aber unbestreitbar wäre die glänzende Entwicklung der Mathematik im 18. Jahrhundert ohne die von Leibniz erfundene *Zeichensprache* nicht möglich gewesen.

Die Literatur über Differential- und Integralrechnung ist außerordentlich umfangreich. Von großer historischer Bedeutung und für die Entwicklung der Infinitesimalrechnung grundlegend sind: Euler, *Institutiones calculi differentialis*, 1755; *Institutiones calculi integralis*, 2. Ausg., 4 Bde., 1792—1794. Lagrange, *Leçons sur le calcul des fonctions*, 2. éd. 1806; *Théorie des fonctions analytiques*, 2. éd. 1813. Cauchy, *Cours d'analyse* 1821; *Résumé des leçons données à l'école polyt.* 1823; *Leçons sur le calcul différentiel*, 1829.

Von neueren Lehrbüchern seien genannt: Cesàro, *Elementares Lehrbuch d. algebr. Analysis und Infinitesimalrechnung* 1904. Kowalewski, *Die klass. Probleme d. Analysis* 1910. Serret-Scheffers, *Lehrb. d. Differential- und Integralr.*, 3 Bde. Stegemann-Kiepert, *Grundriß d. Diff.- u. Integralr.*, 2 Bde. Genocchi-Peano, *Differentialr. und Grundzüge d. Integralr.* 1899. Stolz, *Grundzüge d. Different.- und Integralr.*, 3 Bde., 1893/99. Pascal, *Calcolo infinitesimale*, 3 Bde. Jordan, *Cours d'analyse,* 3 Bde., 2. éd., 1893/96. Goursat, *Cours d'analyse*, 2 Bde., 1902.

Zur ersten Einführung sind die folgenden Werke geeignet, die besonders die Bedürfnisse der angewandten Mathematik im Auge haben: Scheffers, *Lehrb. d. höheren Mathematik*, 1905. Burkhardt, *Vorl. über die Elemente der Diff.- und Integralr.*, 1907. Nernst-Schoenflies, *Einf. in die math. Behandl. der Naturw.* Perry, *Höhere Analysis für Ingenieure*, 1910. Greenhill, *Differential- and Integralcalculus,* 1896.

## § 1. Infinitesimale Größen.

Eine Veränderliche, die bei einem bestimmten Grenzprozeß den Grenzwert *Null* hat, heißt *unendlich klein* oder *infinitesimal.* Zwei unendlich kleine Größen $\alpha$ und $\beta$ heißen *von derselben Ordnung,* wenn $\lim \frac{\alpha}{\beta}$ einen *endlichen, von Null verschiedenen Wert* besitzt. Ist aber $\lim \frac{\alpha}{\beta} = 0$, so nennt man $\alpha$ von *höherer Ordnung,* ist $\lim \frac{\alpha}{\beta} = \infty$, so nennt man $\alpha$ von *niedrigerer Ordnung* als $\beta$.

Ist $\beta$ von höherer Ordnung als $\alpha$ und existiert eine positive Zahl $n$ derart, daß $\lim \frac{\beta}{\alpha^n}$ einen endlichen von Null verschiedenen Wert besitzt, so sagt man, $\beta$ sei *von der Ordnung $n$ in bezug auf $\alpha$.* Nicht immer läßt sich eine derartige Zahl $n$ finden. So ist z. B., wenn $\beta = \frac{1}{\log \alpha}$ ist:

$$\lim \frac{\beta}{\alpha^n} = \begin{matrix} 0 & \text{für } n \leqq 0 \\ \infty & \text{„ } n > 0. \end{matrix}$$

Eine infinitesimale Größe bleibt infinitesimal von derselben Ordnung, wenn sie mit einer endlichen von Null verschiedenen Größe multipliziert wird.

Zwei infinitesimale Größen von gleicher Ordnung unter-

scheiden sich im allgemeinen um eine Infinitesimale derselben Ordnung; nur wenn ihr Verhältnis den *Grenzwert 1* besitzt, unterscheiden sie sich um eine Infinitesimale von *höherer* Ordnung.

*Der Grenzwert des Verhältnisses von zwei infinitesimalen Größen ändert sich nicht, wenn man zu ihnen Infinitesimale höherer Ordnung hinzufügt.*

Die algebraische Summe einer endlichen Anzahl von Infinitesimalen ist wieder infinitesimal, und ihre Ordnung ist mindestens gleich der niedrigsten in der Summe vorkommenden Ordnung.

Das Produkt von zwei infinitesimalen Größen $\alpha$ und $\beta$, die in bezug auf eine dritte Infinitesimale $\gamma$ von den Ordnungen $m$ und $n$ sind, ist infinitesimal von der Ordnung $m + n$ in bezug auf $\gamma$.

## § 2. **Begriff und Existenz des Differentialquotienten.**

Wenn $y = f(x)$ eine Funktion der reellen Veränderlichen $x$ ist und man dem $x$ einen beliebigen Zuwachs $\Delta x = h$ gibt, so erfährt im allgemeinen auch $y$ einen Zuwachs, der $\Delta y$ sei.

Die Grenze des Verhältnisses $\frac{\Delta y}{\Delta x}$ für $\Delta x = 0$ oder

$$\lim_{h=0} \frac{f(x+h) - f(x)}{h}$$

heißt, wenn sie existiert und unabhängig ist von der Art und Weise, in der $\Delta x$ nach Null konvergiert, die *Derivierte* (*Ableitung*) *der Funktion im Punkt x.* Sie wird durch $f'(x)$ oder $y'$ bezeichnet.

Ein *infinitesimaler*, im übrigen aber beliebiger Zuwachs der unabhängigen Variablen $x$ heißt *Differential* von $x$ und wird mit dem Symbol $dx$ bezeichnet.

*Differential der Funktion y von x* heißt das Produkt aus der Ableitung von $y$ mit dem Differential der unabhängigen Variablen (Cauchy, *Exerc. d'anal. et de phys.* **3**, 7 ff. (1844)). Man bezeichnet es mit $dy$, und es ist also

$$dy = f'(x) \cdot dx.$$

Man schreibt deshalb auch die Derivierte

$$f'(x) = \frac{dy}{dx}$$

und nennt sie den *Differentialquotienten* der Funktion $y$ nach $x$. Ist $f'(x)$ stetig, so *unterscheidet sich das Differential der Funktion um Unendlichkleine höherer Ordnung von dem Zuwachs, den die*

*Funktion erfährt, wenn man der unabhängigen Variablen den Zuwachs dx gibt.*

Es kann vorkommen, daß das Verhältnis $\frac{\Delta y}{\Delta x}$ verschiedene Grenzwerte besitzt, je nachdem $\Delta x$ von der *positiven* oder *negativen* Seite her die Grenze *null* erreicht; alsdann spricht man von der Derivierten zur *Rechten* und der zur *Linken* von $x$. (*Vorwärts* und *rückwärts genommener Differentialquotient.*) Unterscheidet man noch genauer zwischen *oberer* und *unterer* Grenze (vgl. Kap. I, § 10), so hat man *vier* Derivierte zu betrachten. Nur wenn diese vier Grenzwerte einander gleich sind, nennt man die Funktion *differenzierbar.*

Allgemeine *notwendige* und *hinreichende* Kriterien für die Differenzierbarkeit einer Funktion lassen sich nicht angeben. Sie ist, wenn sie vorhanden ist, ebenso eine fundamentale Eigenschaft der Funktion wie etwa die Stetigkeit. *Notwendige* Bedingung für die Existenz einer Ableitung in einem Punkt $x$ ist jedenfalls, daß die Funktion in diesem Punkte *stetig* und in ihm und seiner Umgebung *endlich* ist, aber es ist nicht, wie man früher glaubte, jede stetige Funktion auch differenzierbar. Ein von Ampère, *J. éc. pol.* **13**, 148 (1806) versuchter Beweis dieser Behauptung ist nicht stichhaltig (vgl. Du Bois-Reymond, *Journ. f. Math.* **79**, 28 (1875)). Das erste Beispiel einer stetigen und trotzdem in unendlich vielen Punkten *nicht differenzierbaren* Funktion hat Riemann in seiner 1854 verfaßten, aber erst 1867 veröffentlichten Habilitationsschrift gegeben. (*Über die Darstellbarkeit einer Funktion durch eine trigonometrische Reihe. Werke* 2. Aufl. S. 266.) Diese Funktion entsteht durch Integration der Funktion

$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2},$$

wobei das Zeichen $(z)$, sobald $z$ in der Mitte zwischen zwei ganzen Zahlen liegt, den Wert *null*, sonst aber den Überschuß von $z$ über die *nächste* ganze Zahl bedeutet. Das von Riemann angegebene Beispiel führte Hankel zu dem sogenannten *Kondensationsprinzip*, mit dessen Hilfe man, wenn eine Funktion mit einer Singularität in einem Punkt gegeben ist, eine andere Funktion bilden kann, welche dieselbe Singularität in unendlich vielen Punkten besitzt. (Hankel, *Untersuchungen üb. die unendlich oft oszillierenden und unstetigen Funktionen,* Tübingen 1870; *Math. Ann.* **20**, 63 (1882).) Auf Grund dieses Prinzips

lassen sich Funktionen bilden, die in unendlich vielen (aber nicht in allen) Punkten eines Intervalls keine Derivierte haben. Vgl. Dini, *Theorie d. Funkt. einer reellen Größe*, Leipzig 1892, S. 157 ff. Einfacher und weiter tragend ist ein von G. Cantor (*Math. Ann.* **19**, 588 (1882)) angegebenes Kondensationsprinzip. (Dini, a. a. O. S. 188 ff.) Viele Literaturangaben bei Jourdain, *The development of the theory of transfinite numbers, Arch. f. Math. u. Phys.* (3) **10** (1906).

Das bekannteste Beispiel einer stetigen *nirgends differenzierbaren* Funktion gab Weierstraß in seinen Vorlesungen; es wurde zuerst veröffentlicht von Du Bois-Reymond, *Journ. f. Math.* **79**, 29 (1875). Vgl. Wiener, *Journ. f. Math.* **90**, 221 (1881), Weierstraß, *Funktionenlehre* 1886, S. 100, und die besonders anschauliche Darstellung bei F. Klein, *Anw. d. Diff. u. Integralr. auf Geometrie. Autographiertes Vorlesungsheft* 1901. Die Funktion lautet:

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x),$$

worin $a$ zwischen 0 und 1, $b$ eine ungrade ganze Zahl und $ab > 1 + \frac{3}{2}\pi$ ist. Sehr allgemeine Funktionen von derselben Eigenschaft, die die Weierstraßsche Funktion als speziellen Fall enthalten, hat Dini (a. a. O. S. 218 ff.) aufgestellt. Vgl. Lerch, *Journ. f. Math.* **103**, 126 (1888). Von weiteren Arbeiten über stetige nicht differenzierbare Funktionen seien genannt H. A. Schwarz, *Ges. Abhandl.* **2**, 269, Darboux, *Ann. éc. norm.* (2), **4**, 57 (1875), Steinitz, *Math. Ann.* **52**, 58 (1899), H. v. Koch, *Acta math.* **30**, 145 (1907), Landsberg, *Math. Ver.* **17**, 46 (1908), Faber, *Math. Ann.* **66**, 81 (1908). Köpcke (*Math. Ann.* **34**, 161 (1889) und **35**, 104 (1890)) hat gezeigt, daß eine Funktion in jedem endlichen Intervall unendlich viele Oszillationen haben und doch differenzierbar sein kann. Vgl. Schoenflies, *Math. Ann.* **54**, 553 (1901) und desselben Verf. *Bericht üb. die Mengenlehre, Math. Ver.* **8**, 161 ff. Man sieht daraus, daß die Eigenschaft der *Stetigkeit* und *Differenzierbarkeit* noch nicht ausreicht, um eine Funktion $y = f(x)$ als eine *gewöhnliche* (*reguläre*) Funktion zu kennzeichnen, d. h. eine solche, die sich geometrisch durch eine gewöhnliche kontinuierliche *Kurve* veranschaulichen läßt, wenn man $x, y$ als rechtwinklige Punktkoordinaten deutet. Eine derartige Funktion muß nämlich jedenfalls auch *abteilungsweise monoton* sein, d. h. sie darf in jedem endlichen Intervall nur eine endliche Anzahl

Maxima und Minima besitzen. Diese Eigenschaft machte sich zum ersten Mal geltend in der Arbeit von Dirichlet über die Theorie der Fourierschen Reihe (*Journ. f. Math.* **4** (1829), *Werke*, **1**, 131. Vgl. Du Bois-Reymond, *Journ. f. Math.* **79**, 33 (1875)). Über die sonstigen Bedingungen, die eine reguläre Funktion erfüllen muß, vgl. etwa die erwähnte Vorlesung von F. Klein. Entsprechend dem geometrischen Ursprung des Funktionsbegriffes hat man früher (bis zur Aufstellung des allgemeinsten Dirichletschen Funktionsbegriffs (Kap. I, § 9, vgl. Hankel, *Math. Ann.* **20**, 67 (1870)) einer „stetigen" Funktion stillschweigend die Eigenschaften einer *regulären* Funktion zugeschrieben, und es soll auch im folgenden, wo nichts anderes ausdrücklich bemerkt ist, unter einer „Funktion" schlechthin immer eine reguläre Funktion verstanden werden. Wir können dann diesen Paragraphen mit dem folgenden Satz schließen: *Stellt man die Funktion $y = f(x)$ geometrisch durch eine Kurve dar, so gibt der Differentialquotient $\frac{dy}{dx} = f'(x)$ die trigonometrische Tangente des Winkels an, den die Kurventangente im Punkte $(x, y)$ mit der positiven Richtung der x-Achse bildet.*

## § 3. Grundregeln für die Differentiation entwickelter Funktionen einer Veränderlichen.

1. *Der Differentialquotient einer Konstanten ist Null.*

2. *Der Differentialquotient einer algebraischen Summe von Funktionen ist gleich der algebraischen Summe der Differentialquotienten der einzelnen Funktionen.*

Eine in einem Intervall konvergente *unendliche Reihe von Funktionen*

$$F(x) = f_1(x) + f_2(x) + f_3(x) + \cdots$$

heißt *gliedweise differenzierbar,* wenn die aus den Derivierten der einzelnen Funktionen gebildete Reihe konvergiert und als Summe die Derivierte von $F(x)$ ergibt. Es gilt der Satz:

3. *Eine konvergente unendliche Reihe von Funktionen ist gliedweise differenzierbar, wenn die einzelnen Derivierten stetig sind und die aus ihnen gebildete Reihe gleichmäßig konvergiert.*

Insbesondere läßt sich eine *Potenzreihe* für jeden Wert der Variablen, bei dem sie konvergiert, beliebig oft gliedweise differenzieren.

4. *Der Differentialquotient eines Produkts aus einer Konstanten mit einer Funktion ist gleich dem Produkt der Konstanten mit dem Differentialquotienten der Funktion:*

$$\frac{d(ay)}{dx} = a\frac{dy}{dx},$$

*wenn a konstant.*

5. *Ein Produkt von zwei Funktionen wird differentiiert, indem man jede Funktion mit der Derivierten der andern multipliziert und die erhaltenen Produkte addiert:*

$$(uv)' = uv' + vu'.$$

6. *Den Differentialquotienten eines Produktes von mehr als zwei Funktionen findet man am besten nach der folgenden Formel (logarithmische Differentiation):*

$$\frac{(u_1 u_2 \ldots u_n)'}{u_1 u_2 \ldots u_n} = \frac{u_1'}{u_1} + \frac{u_2'}{u_2} + \cdots + \frac{u_n'}{u_n}.$$

7. *Die Derivierte eines Quotienten von zwei Funktionen ist:*

$$\left(\frac{u}{v}\right)' = \frac{vu' - uv'}{v^2}.$$

8. Ist $x = \varphi(y)$ die *inverse Funktion* von $y = f(x)$, so ist *der Differentialquotient von x nach y der reziproke Wert des Differentialquotienten von y nach x:*

$$\varphi'(y) = \frac{1}{f'(x)}.$$

9. Sehr oft kann man eine Funktion $y = F(x)$ nur differentiieren, indem man sie als (einfachere) *Funktion* $f(z)$ *einer anderen Funktion* $z = \varphi(x)$ auffaßt, also $y = f(\varphi(x))$. Man nennt dann $z$ *Zwischenvariable* und hat den Satz:

*Man findet den Differentialquotienten von y nach x, indem man zunächst y nach z differentiiert und mit dem Differentialquotienten der Zwischenvariablen multipliziert:*

$$\frac{dy}{dx} = \frac{dy}{dz} \cdot \frac{dz}{dx}.$$

10. Wir stellen hier die *elementaren Funktionen und ihre Differentialquotienten* zusammen:

$$y = x^n, \qquad y' = n x^{n-1}$$

$$y = e^x, \qquad y' = e^x$$

$$y = a^x, \qquad y' = a^x \ln a$$

$$y = \ln x, \qquad y' = \frac{1}{x}$$

$$y = \sin x, \qquad y' = \cos x$$

$$y = \cos x, \qquad y' = - \sin x$$

$$y = \operatorname{tg} x, \qquad y' = \frac{1}{\cos^2 x}$$

$$y = \operatorname{ctg} x, \qquad y' = - \frac{1}{\sin^2 x}$$

$$y = \arcsin x, \qquad y' = \frac{1}{\sqrt{1 - x^2}} \qquad \begin{matrix} -1 < x < +1 \\ -\frac{\pi}{2} < y < +\frac{\pi}{2} \end{matrix}$$

$$y = \arccos x, \qquad y' = - \frac{1}{\sqrt{1 - x^2}} \qquad \begin{matrix} -1 < x < +1 \\ 0 < y < \pi \end{matrix}$$

$$y = \operatorname{arc\,tg} x, \qquad y' = \frac{1}{1 + x^2}$$

$$y = \operatorname{arc\,ctg} x, \qquad y' = - \frac{1}{1 + x^2}.$$

## § 4. Differentialquotienten höherer Ordnung von Funktionen einer Veränderlichen.

Wiederholt man an der Derivierten von $y$ die Operation des Differentiierens und fährt so fort (vorausgesetzt, daß man immer differentiierbare Funktionen hat), so erhält man die *Differentialquotienten* 2[ter], 3[ter], ... *Ordnung* oder die *zweiten, dritten, ... Derivierten.* Man schreibt sie

$$y'' = f''(x) = \frac{d^2 y}{dx^2},$$

$$y''' = f'''(x) = \frac{d^3 y}{dx^3},$$

und allgemein

$$y^{(n)} = f^{(n)}(x) = \frac{d^n y}{d x^n}.$$

Es ist also *das* $n^{te}$ *Differential der Funktion* $y$:

$$d^n y = f^{(n)}(x)\,(dx)^n,$$

*unter der Voraussetzung, daß man das erste Differential $dx$ der unabhängigen Veränderlichen $x$ als unabhängig von dieser Veränderlichen annimmt.* (Vgl. Cesàro, *Lehrbuch der algebr. Analysis und Infinitesimalrechnung,* 1904, S. 494.)

Eine Funktion, von der man Differentialquotienten beliebiger Ordnung bilden kann, heißt *unbeschränkt differentiierbar.* (Vgl. das Referat von Pringsheim über die Grundlagen der allgemeinen Funktionenlehre in der *Encykl.* **2**, 23.) Besitzt $f(x)$ an der Stelle $x = a$ Differentialquotienten bis zur $n^{\text{ten}}$ Ordnung von endlicher Größe, so ist $f^{(n)}(a)$ der Grenzwert des Verhältnisses $\frac{\Delta^n f(a)}{h^n}$ für $\lim h = 0$, wenn unter $\Delta^n f(a)$ die $n^{\text{te}}$ Differenz der Funktion $f(x)$ für $x = a$:

$$\Delta^n f(a) = f(a + nh) - \binom{n}{1} f(a + (n-1)h) + \binom{n}{2} f(a + (n-2)h) - \cdots + (-1)^{n-1}\binom{n}{1} f(a + h) + (-1)^n f(a)$$

verstanden wird. Der umgekehrte Satz ist nicht richtig, d. h. es kann der Grenzwert existieren, ohne daß die $n^{\text{te}}$ Derivierte $f^{(n)}(a)$ existiert. (Vgl. Harnack, *Math. Ann.* **23**, 260 (1884), Stolz, *Grundz. d. Diff.- u. Int.-Rechnung,* Leipzig 1893, **1**, 93.)

Sind $x$ und $y$ *Funktionen einer unabhängigen Veränderlichen* $t$ und werden ihre Derivierten nach $t$ mit $x'$, $x''$, ..., $y'$, $y''$, ... bezeichnet, so sind die *Derivierten von* $y$ *nach* $x$:

$$\frac{dy}{dx} = \frac{y'}{x'}, \quad \frac{d^2 y}{dx^2} = \frac{x'y'' - y'x''}{x'^3},$$

$$\frac{d^3 y}{dx^3} = \frac{x'(x'y''' - y'x''') - 3x''(x'y'' - y'x'')}{x'^5}.$$

Schon Leibniz hat für die $n^{te}$ *Derivierte eines Produkts von zwei Funktionen* die Formel

$$(uv)^{(n)} = u^{(n)}v + \binom{n}{1} u^{(n-1)} v' + \binom{n}{2} u^{(n-2)} v'' + \cdots + uv^{(n)}$$

gegeben.

Über die allgemeine Theorie der höheren Differentialquotienten und ihre Literatur vgl. den Artikel von Voß in der *Enzykl.* **2**, 87.

Als Beispiele für höhere Differentialquotienten mögen folgende angeführt werden:

$$\frac{d^n(x^m)}{dx^n} = m(m-1)\ldots(m-n+1)x^{m-n};$$

$$\frac{d^n \sin x}{dx^n} = \sin\left(x + \frac{n\pi}{2}\right), \quad \frac{d^n \cos x}{dx^n} = \cos\left(x + \frac{n\pi}{2}\right).$$

$$\frac{d^n}{dx^n}(e^{ax}\cos bx) = r^n e^{ax}\cos(bx + n\vartheta),$$

$$\frac{d^n}{dx^n}(e^{ax}\sin bx) = r^n e^{ax}\sin(bx + n\vartheta),$$

worin $r = \sqrt{a^2 + b^2}$ und $\vartheta = \operatorname{arc\,tg}\frac{b}{a}$ ist.

Ist $f(x) = ax^2 + 2bx + c$ eine quadratische Funktion mit der Diskriminante $\varDelta = b^2 - ac$ und wird $\frac{a(ax^2 + 2bx + c)}{(ax + b)^2} = u$, $\frac{\varDelta}{(ax + b)^2} = 1 - u = v$ gesetzt, so ist

$$\frac{d^n}{dx^n}f(x)^\mu = 2\mu(2\mu-2)\ldots(2\mu-2n+2)\frac{(ax+b)^n}{f(x)^{n-\mu}}F\left(-\frac{n}{2}, -\frac{n-1}{2}, \mu-n+ \right.$$

$$= 2\mu(2\mu-1)\ldots(2\mu-n+1)\frac{(ax+b)^n}{f(x)^{n-\mu}}F\left(-\frac{n}{2}, -\frac{n-1}{2}, \frac{1}{2}-\mu, v\right.$$

wobei die $F$ abbrechende hypergeometrische Reihen in der Bezeichnung von Gauß bedeuten. Die erste Formel bei Schlömilch, *Übungsb.* 5. Aufl. Leipzig 1904, **1**, 52; die zweite für $\mu = -\frac{1}{2}$ bei Hermite, *Cours d'analyse* 1873, p. 310, für dasselbe $\mu$ und $f(x) = 1 - x^2$ schon bei Euler, *Inst. calc. diff.* 1755, p. 150.

Für $n = \varkappa - 1$, $\mu = \varkappa - \frac{1}{2}$ erhält man

$$\frac{d^{\varkappa-1}}{dx^{\varkappa-1}}f(x)^{\varkappa-\frac{1}{2}} = \frac{1\cdot 3\cdot 5\cdots 2\varkappa-1}{a^{\frac{1}{2}}\cdot\varkappa}(ax+b)^\varkappa\cdot\frac{(1+\sqrt{u})^\varkappa - (1-\sqrt{u})^\varkappa}{2},$$

und dies geht für $f(x) = 1 - x^2$ in eine Formel von Jacobi über:

$$\frac{d^{\varkappa-1}(1-x^2)^{\varkappa-\frac{1}{2}}}{dx^{\varkappa-1}} = (-1)^{\varkappa-1}\frac{1\cdot 3\cdot 5\cdots 2\varkappa-1}{\varkappa}\sin(\varkappa\arccos x).$$

Für $\mu = -\frac{1}{2}$ bzw. $\mu = n$ liefert die erste bzw. zweite der obigen Formeln:

$$\frac{d^n}{dx^n}\left(\frac{1}{\sqrt{f(x)}}\right) = (-1)^n n!\,\frac{a^{\frac{n}{2}}}{f(x)^{\frac{n+1}{2}}}P_n\left(\frac{1}{\sqrt{u}}\right),$$

$$\frac{d^n}{dx^n}f(x)^n = 2\cdot 4\cdot 6\cdots 2n\cdot\varDelta^{\frac{n}{2}}P_n\left(\frac{1}{\sqrt{v}}\right),$$

worin $P_n$ die (auch oft mit $X_n$ bezeichnete) *einfache Kugelfunktion,*

auch *Legendresches Polynom* genannt, bedeutet. Zur ersten Formel vgl. Thomson u. Tait, *Natural philos.*, deutsche Ausg. 1871, **1**, 175. Die letzte Formel ist für $f(x) = x^2 - 1$, also $\Delta = 1$, $\frac{1}{\sqrt{v}} = x$ von Rodrigues, *Corr. d. l'éc. polyt.* **3** (1816), 375 gefunden worden. Über Verallgemeinerungen dieser Formel durch Jacobi, Heine, Hermite u. a. vgl. Burkhardt, *Bericht über Entwickl. nach oszill. Funktionen, Math.-Ver.* **10**, § 84 und 85.

Über die Ausdehnung des Begriffs der höheren Differentialquotienten auf nicht ganzzahlige $n$ (*Differentiationen zu beliebigem Index*), die schon von Leibniz, Joh. Bernoulli und Euler ins Auge gefaßt wurde, vgl. Liouville, *J. éc. polyt.* **21**, **24**, **25** (1832/36), Riemann, *Werke*, S. 332, Lindner, *Berl. Math. Ges.* **7**, 77 (1908) (*Arch. Math. u. Phys.* (3) **13** (1908)).

## § 5. Differentiation von Funktionen mehrerer Veränderlichen, von zusammengesetzten und unentwickelten Funktionen einer Veränderlichen.

Eine Funktion $f(x, y, z, \ldots)$ von mehreren Veränderlichen kann man hinsichtlich jeder einzelnen Variablen differentiieren, indem man die anderen Variablen als konstant ansieht. Man erhält dann die *partiellen Derivierten* (Differentialquotienten) erster Ordnung nach $x, y, z, \ldots$ und bezeichnet sie durch $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \ldots$ oder $f_x, f_y, f_z, \ldots$ Weitere Differentiationen liefern die partiellen Derivierten zweiter Ordnung $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial y^2}, \ldots$ oder $f_{xx}, f_{xy}, f_{yx}, f_{yy}, \ldots$ usw. Es besteht der fundamentale Satz von der *Umkehrbarkeit der Differentiationsordnung*, daß nämlich

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

ist. Hinreichende Voraussetzungen dabei sind, daß die beiden partiellen Ableitungen 1. Ordnung und eine Ableitung 2. Ordnung vorhanden und stetig sind; es kann auch für *eine* Ableitung 1. Ordnung die Voraussetzung der Stetigkeit entbehrt werden. (H. A. Schwarz, *Verh. d. Schweiz. naturf. Gesellsch.* 1873, 239 od. *Ges. Abh.* **2**, 275; Timpe, *Math. Ann.* **65**, 310 (1908), woselbst auch weitere Literatur angegeben.)

Infolge dieses Satzes kann man die partiellen Ableitungen $n^{\text{ter}}$ Ordnung mit $\frac{\partial^n f}{\partial x^\alpha \partial y^\beta \partial z^\gamma \dots}$ $(\alpha+\beta+\gamma+\dots=n)$ bezeichnen. Eine Funktion von $\varkappa$ Veränderlichen hat $\frac{\varkappa(\varkappa+1)\cdots(\varkappa+n-1)}{1\cdot 2\cdots n}$ Ableitungen $n^{\text{ter}}$ Ordnung

Die partiellen Derivierten einer *homogenen Funktion*

$$f(x_1, x_2, x_3, \dots)$$

von $n$ Veränderlichen und der Dimension $m$ sind ebenfalls homogen von der Dimension $m-1$, und es ist nach Euler (*Calc. diff.* (1755), 190)

$$x_1\frac{\partial f}{\partial x_1}+x_2\frac{\partial f}{\partial x_2}+x_3\frac{\partial f}{\partial x_3}+\dots+x_n\frac{\partial f}{\partial x_n}=mf.$$

Umgekehrt ist diese Gleichung notwendige und hinreichende Bedingung, daß die Funktion $f$ homogen ist, und $m$ ist dann ihre Dimension. Derselbe Satz liefert für die zweiten Derivierten die Beziehung

$$\sum_{i=1}^{n}\sum_{\varkappa=1}^{n} x_i x_\varkappa \frac{\partial^2 f}{\partial x_i \partial x_\varkappa}=m(m-1)f, \quad \text{usf.}$$

Sind $dx, dy, dz, \dots$ Differentiale der Variablen $x, y, z, \dots$, so sind

$$\partial_x f = f_x dx, \quad \partial_y f = f_y dy, \quad \partial_z f = f_z dz, \quad \dots$$

die *partiellen Differentiale* der Funktion nach den einzelnen Veränderlichen. *Ihre Summe heißt das totale Differential* $df$ *der Funktion,* also

$$df=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dz+\cdots$$

Dieser Ausdruck bleibt gültig, auch wenn die Veränderlichen $x, y, z, \dots$ nicht alle voneinander unabhängig sind, d. h. sind $x, y, z, \dots$ ihrerseits Funktionen von $u, v, w, \dots$, so ist

$$\frac{\partial f}{\partial u}du+\frac{\partial f}{\partial v}dv+\frac{\partial f}{\partial w}dw+\cdots=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dz+\cdots$$

*Sind die ersten partiellen Derivierten stetige Funktionen, so unterscheidet sich das totale Differential um Unendlichkleine höherer Ordnung von dem Zuwachs, den die Funktion erfährt, wenn man die Variablen* $x, y, z, \dots$ *um* $dx, dy, dz, \dots$ *ändert.*

Betrachtet man $dx, dy, dz, \dots$ als *konstant,* so wird das totale Differential von $df$ als *totales Differential zweiter Ordnung*

von $f$ bezeichnet usf. Das totale Differential $n^{\text{ter}}$ Ordnung erhält man durch die symbolische Formel

$$d^n f = [f_x dx]^n.$$

Darin ist $[f_x dx] = f_x dx + f_y dy + f_z dz + \cdots$, und nach erfolgter Potenzierung dieses Ausdrucks ist jedes Produkt

$$f_x^{\alpha} f_y^{\beta} f_z^{\gamma} \ldots \quad (\alpha + \beta + \gamma + \cdots = n)$$

durch die entsprechende Derivierte $n^{\text{ter}}$ Ordnung $\dfrac{\partial^n f}{\partial x^{\alpha} \partial y^{\beta} \partial z^{\gamma} \ldots}$ zu ersetzen. Diese Formel setzt aber für $n > 1$ die *Unabhängigkeit* der $x, y, z, \ldots$ voraus.

Sind $x, y, z, \ldots$ Funktionen einer Veränderlichen $t$, so heißt $f(x, y, z, \ldots)$ eine aus $x, y, z, \ldots$ *zusammengesetzte Funktion* von $t$. Man kann die Differentialquotienten der Funktion $f$ nach $t$ aus den partiellen Derivierten von $f$ und den Differentialquotienten der $x, y, z, \ldots$ nach $t$ zusammensetzen. Es ist

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} + \cdots = \left[\frac{\partial f}{\partial x}\frac{dx}{dt}\right].$$

Die zweite Derivierte ist in symbolischer Schreibweise

$$\frac{d^2 f}{dt^2} = \left[\frac{\partial f}{\partial x}\frac{dx}{dt}\right]^2 + \left[\frac{\partial f}{\partial x}\frac{d^2 x}{dt^2}\right],$$

worin das erste Glied nach dem Potenzieren so, wie eben angegeben, zu behandeln und

$$\left[\frac{\partial f}{\partial x}\frac{d^2 x}{dt^2}\right] = \frac{\partial f}{dx}\frac{d^2 x}{dt^2} + \frac{\partial f}{\partial y}\frac{d^2 y}{dt^2} + \cdots$$

ist.

Sind $x, y, z, \ldots$ unabhängige Funktionen von ebenso vielen anderen Variablen $u, v, w, \ldots$, so ist die Funktionaldeterminante

$$J = \frac{\partial(xyz\ldots)}{\partial(uvw\ldots)}$$

von Null verschieden (vgl. S. 157), und die partiellen Derivierten einer Funktion $f(x, y, z\ldots)$ nach den $x, y, z\ldots$ werden durch die nach $u, v, w\ldots$ in folgender Weise ausgedrückt:

$$\frac{\partial f}{\partial x} = \frac{1}{J}\frac{\partial(fyz\ldots)}{\partial(uvw\ldots)}, \quad \frac{\partial f}{\partial y} = \frac{1}{J}\frac{\partial(xfz\ldots)}{\partial(uvw\ldots)}, \quad \frac{\partial f}{dz} = \frac{1}{J}\frac{\partial(xyf\ldots)}{\partial(uvw\ldots)} \cdots$$

Von besonderer Wichtigkeit für Geometrie und mathematische Physik sind gewisse Verbindungen der partiellen Derivierten

einer Funktion, die gegenüber bestimmten Transformationen *invariant* sind. Es seien von ihnen die *Differentialparameter 1. und 2. Ordnung* einer Funktion von 3 Variablen hervorgehoben:

$$\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2 \quad \text{und} \quad \frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2},$$

die bei jeder Koordinatentransformation ungeändert bleiben. Vgl. Lamé, *J. éc. polyt.* **23**, 215 (1834) und seine späteren Lehrbücher z. B. *Leçons s. l. coord. curvilignes* 1859. Beltrami, *Mem. Acc. Bologna* (2) **8** (1869). Die ebenfalls bereits von Lamé gegebene Transformation dieser Differentialausdrücke auf *krummlinige* Koordinatensysteme hat Jacobi (*J. f. Math.* **36**, 117 (1848)) durch Benutzung des Gaußschen Satzes über mehrfache Integrale (vgl. Kap. VIII, § 7) sehr vereinfacht. Vgl. Riemann-Weber, *Part. Differentialgl.*, Braunschweig (1900) **1**, 94. Systematisch tritt der Gesichtspunkt der Invarianz in den Vordergrund in der *absoluten Differentialrechnung* von Ricci und Levi-Civita, *Math. Ann.* **54**, 125 (1901). Vgl. hierzu Kap. XII, § 3.

Durch eine Gleichung $f(x, y) = 0$ ist $y$ *implicit* als Funktion von $x$ gegeben und heißt *unentwickelte Funktion.* Über den Beweis der Existenz der Funktion $y$ und ihrer Derivierten, der für eine ausgedehnte Klasse von Funktionen $f(x, y)$ zuerst von Cauchy gegeben wurde, vgl. etwa Osgood, *Funktionentheorie* Leipzig 1907 (Sammlung Teubner **20**) **1**, 48. Man findet, wenn $f_y$ von Null verschieden ist:

$$\frac{dy}{dx} = -\frac{f_x}{f_y}, \quad \frac{d^2y}{dx^2} = -\frac{f_{xx}f_y^2 - 2f_{xy}f_xf_y + f_{yy}f_x^2}{f_y^3}.$$

Durch $m$ Gleichungen zwischen $m+1$ Veränderlichen $x$, $y_1$, $y_2, \dots y_m$:

$$f_i(x, y_1, y_2, \dots y_m) = 0 \qquad i = 1, 2, \dots m,$$

sind $y_1, y_2, \dots y_m$ als unentwickelte Funktionen von $x$ definiert, *sobald die Funktionaldeterminante* $J = \frac{\partial(f_1 f_2 \cdots f_m)}{\partial(y_1 y_2 \cdots y_m)}$ *nicht verschwindet* (vgl. S. 157). Ihre Differentialquotienten nach $x$ sind dann

$$\frac{dy_1}{dx} = -\frac{1}{J}\frac{\partial(f_1 f_2 \dots f_m)}{\partial(x y_2 \dots y_m)}, \quad \frac{dy_2}{dx} = -\frac{1}{J}\frac{\partial(f_1 f_2 \dots f_m)}{\partial(y_1 x \dots y_m)}, \dots.$$

## § 6. Rollesches Theorem. Mittelwertsätze. Taylorscher Lehrsatz.

Im folgenden wird eine Funktion für solche Intervalle $(a, b)$ der Variablen betrachtet, innerhalb deren *mit Einschluß der Grenzen* die Funktion *stetig* bleibt.

*Wenn eine Funktion einer Variablen an den Grenzen eines Intervalls gleiche Werte annimmt und innerhalb des Intervalls überall eine Derivierte besitzt, so verschwindet die Derivierte mindestens in einem Punkt innerhalb des Intervalls. Theorem* von Rolle (für ganze rationale Funktionen). Vgl. S. 338.

Es folgt daraus unmittelbar der *Mittelwertsatz*: (Lagrange, *Théor. d. fonct. anal.* 1797; 2. Aufl. 1813 § 39) *Besitzt* $f(x)$ *für alle Werte eines Intervalls* $(a, b)$ *eine Derivierte, so gibt es mindestens einen Wert* $\xi$ *der Variablen innerhalb des Intervalls, so daß*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi)$$

*ist.* Geometrisch besagt dieser Satz, daß es in jedem endlichen Stück einer stetigen Kurve zu der die Endpunkte verbindenden Sehne mindestens eine parallele Tangente gibt.

Mit $b - a = h$ erhält der Satz die Form: *Es gibt eine Zahl* $\vartheta$ *zwischen* 0 *und* 1, *so daß*

$$f(a + h) - f(a) = hf'(a + \vartheta h). \quad 0 < \vartheta < 1.$$

Allgemeiner, aber ebenfalls eine direkte Folge des Rolleschen Theorems ist der Satz (Genocchi-Peano, *Differentialr.* (1899), 317; der entsprechende Satz für $n + 1$ Funktionen und ihre Derivierten bis zur $n - 1^{\text{ten}}$ Ordnung ebenda S. 324):

Sind die Funktionen $f(x)$, $\varphi(x)$, $\psi(x)$ in jedem Punkte eines Intervalls differentiierbar, so gibt es mindestens einen Wert $\xi$ innerhalb des Intervalls, so daß

$$\begin{vmatrix} f'(\xi) & \varphi'(\xi) & \psi'(\xi) \\ f(a) & \varphi(a) & \psi(a) \\ f(b) & \varphi(b) & \psi(b) \end{vmatrix} = 0$$

ist. Für $\psi(\xi) = 1$ folgt daraus der Satz von Cauchy (*Calc. diff.* 1829, S. 37):

$$\frac{f(b) - f(a)}{\varphi(b) - \varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)} \quad \text{oder} \quad \frac{f(a+h) - f(a)}{\varphi(a+h) - \varphi(a)} = \frac{f'(a + \vartheta h)}{\varphi'(a + \vartheta h)}, \quad 0 < \vartheta < 1,$$

und dies liefert für $\varphi(x) = x$ den Mittelwertsatz von Lagrange.

Für eine *Funktion von mehreren Veränderlichen* lautet der Mittelwertsatz: *Besitzt eine Funktion* $f(x_1, x_2, x_3, \ldots)$, *deren Variablen der Reihe nach den Intervallen von* $a_1$ *bis* $a_1 + h_1$, *von* $a_2$ *bis* $a_2 + h_2$, *von* $a_3$ *bis* $a_3 + h_3$ *usw. angehören, für alle Werte der Variablen innerhalb dieser Intervalle sämtliche partiellen Derivierten 1. Ordnung, so gibt es eine Zahl* $\vartheta$ *zwischen* 0 *und* 1, *so daß*

$$f(a_1 + h_1, a_2 + h_2, \ldots) - f(a_1, a_2, \ldots)$$
$$= \left(h_1 \frac{\partial f}{\partial x_1} + h_2 \frac{\partial f}{\partial x_2} + \cdots\right)_{x_1 = a_1 + \vartheta h_1,\; x_2 = a_2 + \vartheta h_2, \ldots}$$

Sehr wichtige Folgerungen aus dem Mittelwertsatz sind die Sätze:

*Eine Funktion, deren Derivierte innerhalb eines Intervalls* $(a, b)$ *einen konstanten Wert* $k$ *besitzt, verhält sich in dem ganzen Intervall wie eine lineare Funktion, d. h. es ist für* $a \leqq x \leqq b$:

$$f(x) = f(a) + k(x - a).$$

*Wenn die Derivierte einer Funktion innerhalb eines Intervalls beständig Null ist, so besitzt die Funktion im ganzen Intervall einen konstanten Wert.*

Über den entsprechenden Satz für Derivierte 2. Ordnung vgl. H. A. Schwarz, *J. f. Math.* **72**, 141 (1870) und den Anhang von Harnack zum 2. Bd. von Serrets Diff.- u. Int.-Rech.

Der *Taylorsche Satz* (Taylor, *Methodus incrementorum* 1715) ist zunächst eine Verallgemeinerung des Mittelwertsatzes: *Wenn eine Funktion* $f(x)$ *in einem Intervall von* $a$ *bis* $a + h$ *mit Einschluß der Grenzen endlich ist und alle Differentialquotienten bis zur* $n^{ten}$ *Ordnung besitzt, so ist*

$$f(a + h) = f(a) + \frac{h}{1!} f'(a) + \frac{h^2}{2!} f''(a) + \cdots$$
$$+ \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(a) + R_n,$$

*wobei der Rest* $R_n$ *mit Hilfe einer im ganzen Intervall endlichen und differentiierbaren Funktion* $\varphi(x)$, *deren Derivierte im Intervall nicht verschwindet, durch die Gleichung*

$$R_n = \frac{\varphi(a + h) - \varphi(a)}{\varphi'(a + \vartheta h)} \frac{(1 - \vartheta)^{n-1} h^{n-1}}{1 \cdot 2 \cdot 3 \cdots (n-1)} f^{(n)}(a + \vartheta h) \quad 0 < \vartheta < 1$$

*bestimmt ist.*

Diese sehr allgemeine Form des Restes stammt von Schlömilch (*Handb. d. Diff.- u. Integralr.* 1847).

Für $\varphi(x) = (a + h - x)^p$ mit einem beliebigen *positiven* Exponenten $p$ ergibt sich

$$R_n = \frac{h^n(1-\vartheta)^{n-p}}{p(n-1)!} f^{(n)}(a + \vartheta h),$$

und wenn man hier $p = n$ oder $p = 1$ nimmt, erhält man

$R_n = \frac{h^n}{n!} f^{(n)}(a + \vartheta h)$, Lagrange, *Théorie d. fonct. anal.* (1797) § 40;

$R_n = \frac{h^n(1-\vartheta)^{n-1}}{(n-1)!} f^{(n)}(a + \vartheta h)$, Cauchy, *Résumé* (1823), 173.

Am kürzesten wird der Taylorsche Satz mit Hilfe der *partiellen Integration* abgeleitet (Prony 1805 nach Angabe von Cauchy, *Résumé* 1823); dabei erhält man den *Rest* in *exakter* Form, d. h. ohne die unbestimmte Größe $\vartheta$, als *bestimmtes Integral* (Lagrange, *Théor. d. fonct. anal.*):

$$R_n = \frac{1}{(n-1)!}\int_0^h f^{(n)}(x+t)(h-t)^{n-1}dt,$$

woraus sich die oben angeführten Ausdrücke für $R_n$ ableiten lassen.

Der Taylorsche Satz für $a = 0$, $h = x$ heißt *Maclaurinscher Lehrsatz.* Maclaurin, *Treatise of fluxions* 1742, jedoch bereits von Taylor gegeben.

Besitzt die im Intervall $a$ bis $a + h$ stetige Funktion $f(x)$

1. *innerhalb des Intervalls nur endliche Derivierte von beliebig hoher Ordnung*, und ist

2. $\lim_{n=\infty} R_n = 0$ für jedes zwischen 0 und 1 liegende $\vartheta$,

so besteht für $f(a + h)$ die nach Potenzen von $h$ fortschreitende *Taylorsche Entwicklung*

$$f(a + h) = f(a) + hf'(a) + \frac{h^2}{1\cdot 2}f''(a) + \cdots.$$

Diese *beiden* Bedingungen sind für das Bestehen der Entwicklung *hinreichend*, nicht aber — wie Lagrange (*Th. d. fonct.* chap. V, § 30) vermutete — die erste allein. Ebensowenig ist dessen Ansicht richtig, daß die Reihe, sobald sie konvergiere, den Wert $f(a + h)$ ergebe; bereits Cauchy (*Calc. inf.* 1823, 152 u. 299) hat darauf hingewiesen, daß z. B. in

der (reellen) Umgebung von $x = 0$ die Funktionen $f(x)$ und $f(x) + e^{-\frac{1}{x^2}}$ die gleiche Taylorentwicklung besitzen. Überzeugendere Beispiele bei du Bois-Reymond, *Math. Ann.* **21**, 114 (1876) und Pringsheim, *Münch. Ber.* (1892), 211 oder *Math. Ann.* **42**, 153 (1893). Die abschließende Arbeit von Pringsheim, *Math. Ann.* **44**, 57 (1894) gibt die *notwendigen* und *hinreichenden* Bedingungen für die Darstellbarkeit einer Funktion durch die Taylorsche Reihe, wobei die erste Bedingung im wesentlichen bestehen bleibt, während die *zweite* Bedingung durch die weniger umfassende ersetzt werden kann, *daß mit wachsendem n das Cauchysche Restglied für alle in Betracht kommenden Werte von $\vartheta$ gleichmäßig gegen Null konvergiert.* Vgl. hierzu Pascal, *Esercizi e note critiche di calc. infinit.* (1895), 176 ff., W. H. Young, *Quart. Journ.* **40**, 157 (1909).

Eine erschöpfende Behandlung der Taylorschen Reihe erfordert die Heranziehung *komplexer* Werte der Variablen und ist Sache der *Funktionentheorie* (s. d.)

Eine eingehende Darstellung der Geschichte des Taylorschen Satzes hat Pringsheim, *Bibl. Math.* (3) **1**, 433 (1900) gegeben.

Für *Funktionen mehrerer Veränderlichen* lautet der *Taylorsche Satz: Besitzt eine Funktion $f(x_1, x_2, x_3 \ldots)$, deren Variable der Reihe nach den Intervallen $(a_1, a_1 + h_1)$, $(a_2, a_2 + h_2)$, $(a_3, a_3 + h_3)$ usw. angehören, für alle Werte der Variablen innerhalb dieser Intervalle sämtliche Derivierten bis zur $n^{ten}$ Ordnung, so ist in symbolischer Schreibweise*

$$f(a_1 + h_1, a_2 + h_2, \cdots) = f(a_1, a_2, \cdots)$$
$$+ \left[h\frac{\partial f}{\partial a}\right] + \frac{1}{2!}\left[h\frac{\partial f}{\partial a}\right]^2 + \frac{1}{3!}\left[h\frac{\partial f}{\partial a}\right]^3 + \cdots + \frac{1}{(n-1)!}\left[h\frac{\partial f}{\partial a}\right]^{n-1} + R_n.$$

Darin ist $\left[h\frac{\partial f}{\partial a}\right] = h_1\frac{\partial f}{\partial a_1} + h_2\frac{\partial f}{\partial a_2} + \cdots$, und die Potenzen dieses Ausdruckes sind in derselben Weise zu behandeln wie im vorigen Paragraphen; das Restglied $R_n$ kann man in verschiedener Form geben, am einfachsten:

$$R_n = \frac{1}{n!}\left[h\frac{\partial f}{\partial x}\right]^n_{\substack{x_1 = a_1 + \vartheta h_1 \\ x_2 = a_2 + \vartheta h_2 \\ \cdots\cdots}} \qquad (0 < \vartheta < 1).$$

## § 7. Grenzwerte von Ausdrücken in unbestimmter Form.

Eine Funktion von der Form $F(x) = \frac{f(x)}{\varphi(x)}$ erscheint *unbestimmt*, wenn für $x = a$ entweder beide Funktionen $f(x)$ und $\varphi(x)$ *Null* oder beide *unendlich* werden. Unter dem Wert der Funktion $F(x)$ an dieser Stelle versteht man den *Grenzwert*, wenn ein solcher existiert, den der Quotient für $x = a$ besitzt. Die Funktionen $f(x)$ und $\varphi(x)$ mögen in der *Umgebung* von $a$ stetig und differentiierbar sein. Dann gelten die Sätze:

1. *Wenn $f(x)$ und $\varphi(x)$ für $x = a$ stetig sind und dort verschwinden, so ist*

$$\lim_{x=a} \frac{f(x)}{\varphi(x)} = \lim_{x=a} \frac{f'(x)}{\varphi'(x)}.$$

2. *Wenn $f(x)$ und $\varphi(x)$ für $x = a$ unendlich werden,* wenn ferner $\varphi'(x)$ an dieser Stelle von Null verschieden ist und in der Umgebung das Vorzeichen nicht ändert, so ist ebenfalls

$$\lim_{x=a} \frac{f(x)}{\varphi(x)} = \lim_{x=a} \frac{f'(x)}{\varphi'(x)}.$$

Wenn man bei Anwendung dieser Sätze wieder zu Ausdrücken der Form $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ gelangt, so geht man zu höheren Differentialquotienten weiter; es kann aber vorkommen, daß man mit dem Verfahren nicht zu Ende kommt und gleichwohl ein Grenzwert für den Quotienten existiert. (Vgl. Pascal, *Esercizi e note crit.* 240.)

Der erste Satz geht auf Joh. Bernoulli (1694, Brief an l'Hospital), der zweite auf Euler (*Inst. calc. diff.* (1755), 738) zurück. Die strenge Behandlung beginnt mit Cauchy; abschließend bei Stolz, *Math. Ann.* **14**, 231 (1879).

Andere unbestimmte Formen, wie $0 \cdot \infty$, $\infty - \infty$, $0^0$, $\infty^0$, $1^\infty$ können stets auf die oben behandelten zurückgeführt werden. Die folgenden Beispiele zeigen Eigenschaften der darin vorkommenden Funktionen, von denen oft Gebrauch zu machen ist.

$$\lim_{x=0} \frac{\sin x}{x} = 1, \quad \lim_{x=0} \frac{\operatorname{tg} x}{x} = 1,$$

$$\lim_{x=\infty} x^\alpha e^{-mx} = 0 \quad \text{und} \quad \lim_{x=\infty} \frac{(\ln x)^m}{x^\alpha} = 0$$

für jedes positive $\alpha$ und $m$.

Wenn sich die Funktionen $f(x)$ und $\varphi(x)$ in der Umgebung von $a$, d. h. für $x = a + h$ in *Reihen* nach Potenzen von $h$ entwickeln lassen, so ist es oft vorteilhafter, an Stelle des obigen Verfahrens diese Reihenentwicklungen einzusetzen, mit der niedrigsten im Nenner vorkommenden Potenz von $h$ zu kürzen und zur Grenze $h = 0$ überzugehen.

Über unbestimmte Ausdrücke bei Funktionen von *mehreren* Veränderlichen vgl. Genocchi-Peano, *Differentialr.* (1899), 170 ff.

## § 8. Wachsende und abnehmende Funktionen. Maxima und Minima.

Eine Funktion $f(x)$ heißt im Punkte $x$ *wachsend*, wenn sich ein Intervall, zu dem $x$ gehört, angeben läßt, so daß innerhalb des Intervalls mit jedem $x_1 > x$ auch $f(x_1) > f(x)$, dagegen mit jedem $x_0 < x$ auch $f(x_0) < f(x)$ ist; die Funktion heißt *abnehmend* in $x$, wenn innerhalb eines ebensolchen Intervalls mit jedem $x_1 > x$ zugleich $f(x_1) < f(x)$, dagegen mit jedem $x_0 < x$ zugleich $f(x_0) > f(x)$ ist.

*Ist $f'(x)$ positiv, so ist $f(x)$ wachsend, ist $f'(x)$ negativ, so ist $f(x)$ abnehmend.*

Wenn für den Wert $x$ alle Derivierten bis zur $(n-1)^{\text{ten}}$ Ordnung verschwinden und die erste nicht verschwindende Derivierte $f^{(n)}(x)$ *ungerader* Ordnung ist, so ist $f(x)$ wachsend oder abnehmend, je nachdem $f^{(n)}(x)$ positiv oder negativ ist. Ist aber $f^{(n)}(x)$ von *gerader* Ordnung, so nimmt die Funktion im Punkt $x$ weder zu noch ab, was geometrisch durch eine zur $x$-Achse parallele Tangente gekennzeichnet ist.

Eine Funktion $f(x)$ hat im Punkte $x$ ein *Maximum*, wenn die Funktionswerte an den zu $x$ benachbarten Stellen sämtlich *kleiner* als $f(x)$ sind, d. h. wenn sich ein Intervall, *innerhalb* dessen $x$ liegt, angeben läßt, so daß für irgendeinen nicht mit $x$ zusammenfallenden Punkt $x_1$ des Intervalls $f(x_1) < f(x)$ ist. Dagegen hat $f(x)$ in $x$ ein *Minimum*, wenn die Funktionswerte an den zu $x$ benachbarten Stellen sämtlich *größer* als $f(x)$ sind, d. h. wenn für jedes nicht mit $x$ identische $x_1$ des Intervalls $f(x_1) > f(x)$ ist. Als gemeinsamer Name für Maximum und Minimum hat sich die Bezeichnung *Extremum* eingebürgert.

*Soll $f(x)$ im Punkt $x$ ein Extremum haben, so muß dort die erste Derivierte Null und die erste nicht verschwindende*

*Derivierte $f^n(x)$ von gerader Ordnung sein, und zwar wird $f(x)$ ein Maximum oder Minimum, je nachdem $f^{(n)}(x)$ einen negativen oder positiven Wert hat.*

In der Sache übereinstimmend, bei Anwendungen manchmal brauchbarer ist folgendes Kriterium: *Bei einem Extremum von $f(x)$ wechselt $f'(x)$ das Vorzeichen; geht es dabei (bei wachsendem $x$) von positiven zu negativen Werten, so hat $f(x)$ ein Maximum, andernfalls ein Minimum.*

Über die Fälle, in denen dieser Satz versagt, z. B. $f(x)$ nicht differentiierbar, mit einseitigen Differentialquotienten u. a. vgl. Pascal, *Esercizi etc.* 215 ff., Stolz, *Grundz. d. Diff.- u. Int.-Rech.* (1893) **1**, 206.

Eine Funktion von *mehreren Veränderlichen* hat an einer Stelle $(x_1, x_2, \ldots)$ ein *Maximum* oder *Minimum*, wenn der Wert der Funktion an dieser Stelle *größer* oder *kleiner* ist als sämtliche Funktionswerte in der Umgebung; sind also $\alpha_i$, $\beta_i$ $(i = 1, 2, 3 \ldots)$ positive Größen und $h_1, h_2, \ldots$ irgendwelche Zahlen, die den Ungleichungen $-\alpha_i < h_i > \beta_i$ genügen und nicht alle gleichzeitig Null sind, so muß bei einem

$$\textit{Maximum:}\quad f(x_1 + h_1, x_2 + h_2, \ldots) - f(x_1, x_2, \ldots) < 0,$$
$$\textit{Minimum:}\quad f(x_1 + h_1, x_2 + h_2, \ldots) - f(x_1, x_2, \ldots) > 0$$

sein.

Für ein Extremum der Funktion $f(x_1, x_2, \ldots x_n)$ ist *notwendige* Bedingung, *daß die ersten partiellen Differentialquotienten sämtlich verschwinden* und daß die niedrigste Ordnung, bei der nicht alle partiellen Differentialquotienten Null sind, eine *gerade* Zahl ist.

Sei $k$ diese Zahl, so hängt die Entscheidung, ob ein *Maximum* oder *Minimum* stattfindet, von dem Verhalten des im Taylorschen Satz auftretenden symbolischen Ausdrucks

$$\left[h \frac{\partial f}{\partial x}\right]^k = F$$

ab. Dieser stellt eine Form $k^{\text{ten}}$ Grades in $h_1, h_2, \ldots, h_n$ vor. Für solche Formen (mit geradem $k$) gelten die Begriffe *definit*, *indefinit* und *semidefinit*, wie sie auf S. 123 für *quadratische* Formen erläutert sind. Ein Extremum ist sicher vorhanden, wenn die Form *definit* ist, und zwar ein *Maximum*, wenn sie beständig *negativ*, ein *Minimum*, wenn sie beständig *positiv* ist. Ist $F$ *indefinit*, so ist die betreffende Stelle kein Extremum der

Funktion. Ist aber die Form *semidefinit*, d. h. verschwindet sie für Werte $x_1, x_2, \ldots, x_n$, die nicht alle Null sind, während sie für alle übrigen Wertesysteme dasselbe Vorzeichen hat, so erfordert die Entscheidung, ob ein Extremum stattfindet, spezielle Untersuchungen, wegen derer auf Genocchi-Peano, *Differentialr.*, Scheeffer, *Math. Ann.* **35**, 541 (1885), Stolz, *Grundzüge* **1**, 211 ff. verwiesen sei.

Sei nun $k = 2$, also mögen für eine Stelle $(x_1, x_2, \ldots, x_n)$ alle ersten, aber nicht alle zweiten Differentialquotienten Null sein, so ist $F$ die quadratische Form

$$F = \sum_{i=1}^{n} \sum_{\varkappa=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k} h_i h_k = \sum \sum f_{ik} h_i h_k,$$

wenn $\frac{\partial^2 f}{\partial x_i \partial x_k} = f_{ik}$ gesetzt wird. Bildet man die Reihe der Determinanten

$$D_\alpha = \begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1\alpha} \\ f_{21} & f_{22} & \cdots & f_{2\alpha} \\ \vdots & & & \\ f_{\alpha 1} & f_{\alpha 2} & \cdots & f_{\alpha\alpha} \end{vmatrix} \quad \text{für } \alpha = 1, 2, \ldots, n,$$

so ist $F$ *positiv definit*, wenn *alle* $D_\alpha$ *positiv*, dagegen *negativ definit*, wenn die $D_\alpha$ *abwechselnd negativ oder positiv* (mit *negativem* $D_1 = f_{11}$) sind.

Bei einer Funktion von *zwei* Variablen ist $F$ *definit*, wenn $D_2 = \frac{\partial^2 f}{\partial x_1^2} \frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$ *positiv* ist und die Funktion besitzt alsdann ein *Maximum* oder *Minimum*, je nachdem $\frac{\partial^2 f}{\partial x_1^2}$ $\left(\text{und damit auch } \frac{\partial^2 f}{\partial x_2^2}\right)$ *negativ* oder *positiv* ist. (Lagrange, *Misc. Taur.* **1** (1759).) Ist $D_2 < 0$, so ist $F$ *indefinit* und kein Extremum vorhanden; der *semidefinite* Fall ist durch $D_2 = 0$ gekennzeichnet.

Unter einem *bedingten Maximum* oder *Minimum (relative Extreme)* versteht man ein Extremum der Funktion $f(x_1, x_2, \ldots x_n)$, wobei zwischen den Variablen $k < n$ Bedingungsgleichungen $\varphi_1 = 0$, $\varphi_2 = 0, \ldots, \varphi_k = 0$ zu erfüllen sind. Man führt mit Lagrange (*Th. d. fonct.* 268) zweckmäßig *Multiplikatoren* $\lambda_1, \lambda_2, \ldots, \lambda_k$ ein und bildet die $n$ Gleichungen

$$\frac{\partial f}{\partial x_i} + \sum_{\alpha=1}^{k} \lambda_\alpha \frac{\partial \varphi_\alpha}{\partial x_i} = 0 . \qquad (i = 1, 2, \dots n)$$

Diese müssen im Fall eines Extremums erfüllt sein und liefern zusammen mit den gegebenen $k$ Gleichungen Werte für $x_1 \dots x_n$, $\lambda_1 \dots \lambda_k$. Ob aber diesen Lösungen wirklich Extreme entsprechen und welcher Art sie sind, ist durch weitere Untersuchung zu entscheiden. (Vgl. Stolz, *Grundzüge* **1**, 245.) Diese Methode ist in der *Ausgleichungsrechnung* bei der Ausgleichung bedingter Beobachtungen von Wichtigkeit. (Korrelatenmethode. Gauß, *Theor. comb. observ. Suppl.* § 11 (1826).)